International Journal for Quality in Health Care 2007; pp. 1–9 10.1093/intqhc/mzm040

Content of quality-of-life instruments is affected by item-generation methods

ANNE-CHRISTINE RAT[1], JACQUES POUCHOT[2], FRANCIS GUILLEMIN[3], MICHÈLE BAUMANN[4], NATHALIE RETEL-RUDE[5], ELIZABETH SPITZ[6] AND JOËL COSTE[7]

[1]EA4003, Centre d'Epidémiologie Clinique CIC-EC-Inserm CIE6, Service d'Epidémiologie et Evaluation Cliniques, Nancy University, Nancy, France, [2]Service de Médecine Interne, Hôpital Européen Georges Pompidou, Paris, France, [3]EA4003, Centre d'Epidémiologie Clinique CIC-EC-Inserm CIE6, Service d'Epidémiologie et Evaluation Cliniques, Nancy University, Nancy, France, [4]EA4003, Nancy, France, [5]Laboratoire de Biostatistiques Pharmaceutiques, Université de Franche-Comté, Besançon, France, [6]Laboratoire de Psychologie de la Santé, UFR Sciences Humaines et Arts, Université de Metz, Metx Cedex 1, France, and [7]Départment de Biostatistique et d'Informatique Médicale, Hôpital Cochin, Paris, France

Abstract

Background. Methods used to generate items for complex measurement scales are heterogeneous and probably produce heterogeneous data, yet nothing is known about the advantages of one method over another.

Objective. We aimed to compare methods of generating items for tools designed to measure quality-of-life for patients.

Methods. We used five methods to develop a quality-of-life instrument for patients with lower-limb osteoarthritis: individual interviews with patients involving two different techniques (semi-structured and cognitive), individual interviews with health professionals, and focus groups of patients and health professionals. The process generated 80 items, of which 37 were excluded after content and psychometric analysis. With the final 43-item scale used as a 'reference standard', we estimated the contribution of each method.

Results. For health professionals, the focus group and individual interviews produced 35 and 81% of the items, respectively. For patients, the focus groups produced 74% of the items and both interview techniques 100% of the items. Health professionals provided a narrower picture of the effects of the disease on quality-of-life. Focus groups contributed less to social domains than did individual interviews. The two patient interview techniques highlighted different themes.

Conclusion. In developing a complex measurement scale for patients, we found individual interviews with patients the best method for formulating items; other methods such as physician interviews and focus groups contributed no additional information. Reports of instrument generation should include details of the item-generation step, the methods used to develop items and the number of people involved.

Keywords: instrument development, item generation, osteoarthritis, qualitative methods, quality-of-life, quantitative methods

In health care, the multidisciplinary team care aims to preserve and improve patients' health-related quality-of-life by reducing disease activity and improving functional ability, mental and social health and vocational status. Measuring what matters to patients should remain a priority for quality-of-care research and facilitates client centeredness. Moreover, the selection of appropriate outcome measures is critical to influence public health policies.

The increasing interest in measures reflecting patients' viewpoints has led to high demand for reliable and valid standardized quality-of-life questionnaires. The process of developing a new quality-of-life instrument, including the necessary qualitative and quantitative steps, is well established [1, 2]. However, it is long, complex and technically demanding, requiring input from patients, clinicians, psychologists, sociologists, linguists, psychometricians and statisticians.

The first step in the development of a questionnaire is to generate and obtain items relevant to the question and the targeted population. After literature review, interviews are the most widely used source of qualitative data. However, a wide range of interview methods is available for patients or health professionals individually or in groups. Structured interviews involve asking specific questions for more focused information; unstructured interviews, suggesting a topic, which is discussed freely; and semi-structured interviews, suggesting a topic and patients addressing it as they wish, then asking specific questions to elicit more focused information.

Address reprint requests to Anne-Christine Rat, Centre d'Epidémiologie Clinique, CIC-EC-Inserm CIE6, Service d'épidémiologie et évaluation cliniques, hôpital Marin- CHU de Nancy, 2 avenue du Maréchal de Lattre de Tassigny, C.O N°34, 54035 Nancy Cédex, France. Tel: +33-3-83852163; Fax: +33-3-83851205; E-mail: ac.rat@chu-nancy.fr

So far, cognitive interviews have rarely been used during the development of quality-of-life questionnaires [3–9]. In most studies, they have been used to examine the content validity of instruments by investigating how well patients perceived and interpreted questions. Cognitive interviews are based on memory retrieval, knowledge representation and communication [10, 11], to facilitate patients' recall of a situation in the emotional and environmental context, encourage them to report as many details as possible, and narrate events from various points of view [12].

Focus-group participants provide mutual support and share feelings and experience, interactions between individuals producing insights that would not surface in individual interviews. Moreover, data are obtained rapidly, and specific themes or new ideas emerge [13, 14].

Given that questionnaire items should reflect the patient's perspective and be acceptable, direct elicitation of patient experience might provide the broadest and more pertinent data. However, health professionals use their own experience to evaluate patients with various disabilities. They have different priorities [15, 16] and distance from the disease, and may yield valuable items for instruments [17].

Methods used to generate items for complex measurement scales are heterogeneous and probably produce heterogeneous data, yet nothing is known about each method's advantages. In fact, the generation of items has rarely been studied, and few reports of instrument development describe in detail items generation [3–5, 18–21]. To our knowledge, no comparisons of methods used to generate items exist in the literature, and no published guidelines for their use.

This paper compares the relative contribution of several methods of item generation to the development of a quality-of-life instrument for osteoarthritis, the OsteoArthritis of Knee and Hip Quality-Of-Life (OAKHQOL) scale [22].

Patients and methods

Development of the OAKHQOL (Fig. 1)

Item generation

Methods used to elicit oral material. We elicited oral expressions from patients and health professionals in five ways [23]: (i) Semi-structured patient interviews conducted in two phases: spontaneous conversation, then probing to identify further problems. To determine which themes of quality-of-life to explore, we created an interview guide by analysing the content of unstructured interviews from a pilot study of 16 patients not included in the study sample, 16 health professionals and 2 focus groups of patients. The only instruction given to interviewees was to describe the impact of osteoarthritis on their quality-of-life, difficulties they encountered because of osteoarthritis and which events they had trouble with in daily life. No quality-of-life definition was suggested prior to or during the interview. Each semi-structured interview lasted about an hour. Objectivity was guaranteed because interviewers were psychologists and

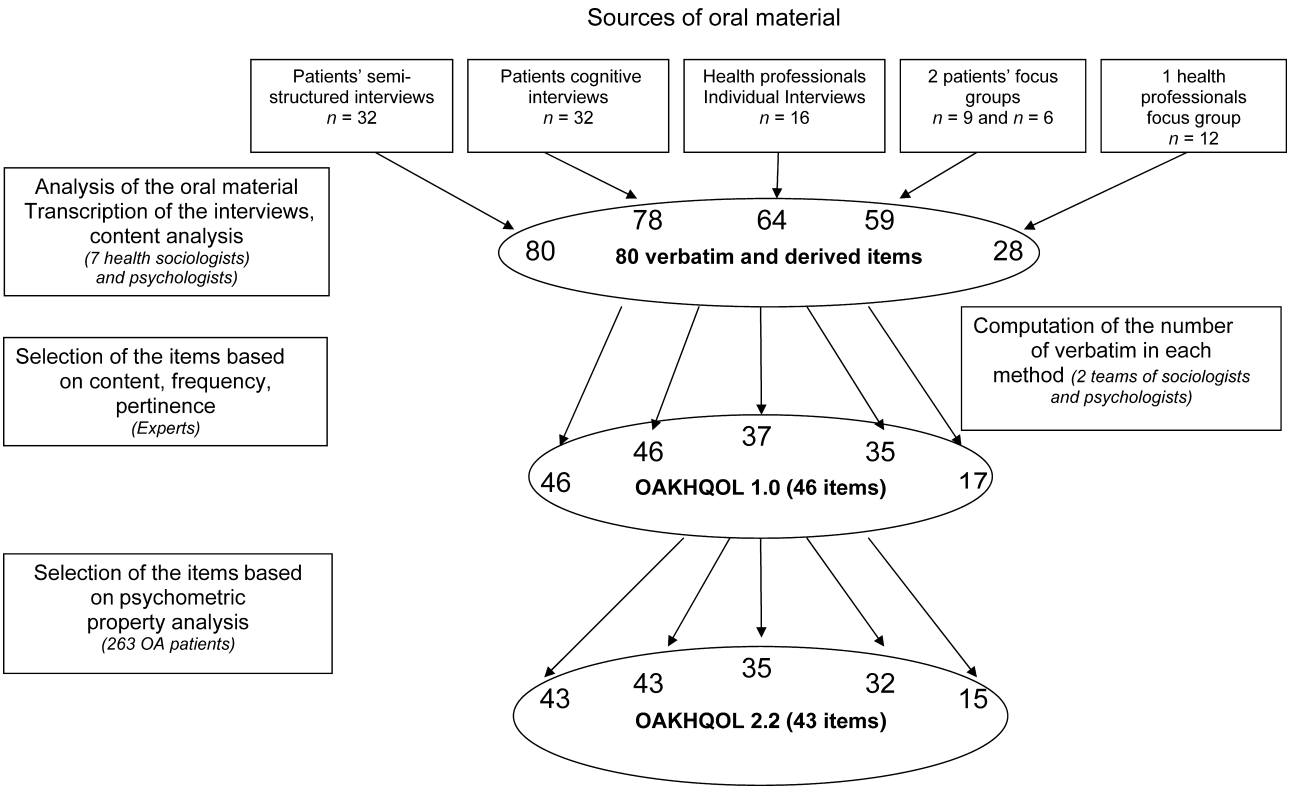

Figure 1 Development of the OAKHQOL.

sociologists trained to conduct interviews but had no knowledge of quality-of-life or osteoarthritis. (ii) Cognitive interviews [10] with an interview guide from 16 exploratory interviews of patients involving a cognitive interview technique. Each interview lasted about an hour, and interviewers were trained to conduct cognitive interviews but had no knowledge of quality-of-life or osteoarthritis. (iii) Unstructured hour-long interviews with health professionals involved in osteoarthritis care. (iv) Unstructured focus-group sessions with patients. (v) Unstructured focus-group sessions with health professionals [24]. Focus-group sessions lasted about 2h and were conducted by a psychologist trained to conduct such groups.

Sample (Fig. 1). Patients with a diagnosis of hip or knee osteoarthritis according to the American College of Rheumatology criteria [25, 26] who had an appointment in rheumatology or in orthopaedic surgery outpatient clinics were asked to participate. To construct the study sample, we used a quota system by age, sex, osteoarthritis location (hip or knee) and medical or surgical stage of osteoarthritis, which would elicit comprehensive and relevant items.

The health professionals surveyed were all familiar with management of osteoarthritis and included general practitioners, rheumatologists, orthopedic surgeons, rehabilitation specialists, physiotherapists, nurses and occupational therapists.

A total of 79 patients and 28 health professionals participated in interviews or focus groups. Sixty-four patients were interviewed in semi-structured and cognitive interviews, and 16 health professionals participated in individual unstructured interviews. The three focus groups had 9, 6 and 12 people, respectively (Fig. 1). Patients and health professionals were recruited for interviews or focus groups in the same way. All participants were told about the aims of the study and objectives of the interviews or focus groups. Oral consent was obtained for all participants, but no institutional ethics review board approval was necessary.

Analysis of oral material. All interviews and focus-group sessions were recorded on tape and transcribed verbatim. Six health sociologists and psychologists working independently in pairs conducted a semantic theme content analysis of a third of the transcripts each. Both analysts grouped the content of the transcripts into categories or themes. The pair would then reach a consensus that was reviewed with a senior academic sociologist. Initial analysis of interview transcripts identified 80 categories.

Source material was then independently analysed by two other teams of psychologists and sociologists to determine which items of the list resulted from the five methods of item generation (Fig. 1).

Selection of items and questionnaire construction. Our approach to item selection was developed by the French Quality-of-Life in Rheumatology group [27, 28]. It combines psychometric and expert information and creates a hierarchy of priorities that favors the content of the item over its psychometric properties. Indeed, an expert-based approach, possibly helped by statistical considerations, appears preferable to statistical approach only [27]. Experts' assessment was carefully organized to ensure impartiality and representativeness and to limit information bias. The 80 items were reviewed by 10 experts not involved in the item-generation step: two osteoarthritis patients, one psychologist, one sociologist, two rheumatologists, one rehabilitation specialist, one orthopaedic surgeon, one linguist and one epidemiologist. The overall concept of the OAKHQOL was based on the World Health Organization definition of quality-of-life [29] and according to the expert panel, some items did not pertain to quality-of-life but, rather, satisfaction with care ($n = 3$), locus of control ($n = 1$) and coping strategies ($n = 13$). Several items referred to the ways of dealing with pain (bearing it, getting used to it, disappointment with pain medication and powerlessness). Others referred to the ways of coping with disability (changes in employment, being practical in everyday life) and attitudes concerning the self (needing to work or to be occupied to think about something else, putting aside the disease and no longer being able to please or like oneself). Finally, issues about medications and treatments that were raised but not selected, included the need for explanations and support from health professionals, preferences for non-drug treatment and declining to take drugs because of lack of improvement or worry about getting used to the drug. Most of these items were suggested by $<50\%$ of patients.

By reference to the transcripts, items included in the questionnaire were further worded to ensure acceptability and understanding for the first version of the OAKHQOL (1.0).

Classical test theory was used to document psychometric properties; analyses of response rates, floor and ceiling effects, dimensionality, construct validity and reliability were conducted at the level of items and dimensions. Finally, on quantitative analysis of the OAKHQOL (1.0), three items were eliminated because of their psychometric properties (low test–retest reliability in two, and low response rate and absence of loading on any factor in one) [22]. Thus, item selection in the OAKHQOL combined assessment of content and psychometric properties, giving greater weight to the former. The final version (2.2) of the OAKHQOL is a self-administered questionnaire of 43 items, grouped into five dimensions: physical activities, mental health, pain, social activities and social support. Three items are independent [22].

Statistical analysis

We considered the final version of the OAKHQOL as the reference or criterion standard for comparing material provided by each method. We compared the proportions ($P+$) of the items of the final version of the OAKHQOL generated by each of the five methods, $P+$ being the number of items generated by the method that were also included in the final questionnaire over the total number of final questionnaire items. The proportion of non-selected items ($P-$) for each method in comparison to the others was also determined, $P-$ being the number of items generated by the considered method and not included in the final questionnaire over the total number of items generated by the all methods and not included in the final questionnaire. These proportions and their confidence intervals were computed for each method and compared by paired chi-square test.

Table 1 Proportion of the items of the final version of the OAKHQOL (P+) and proportion of non-selected items (P−) for each method of generating items

	Patient semi-structured interviews ($n = 32$)		Patient cognitive interviews ($n = 32$)		Health professional individual interviews ($n = 16$)		Patient focus groups ($n = 9$ and 6)		Health professional focus group ($n = 12$)	
	N		*N*		*N*		*N*		*N*	
Items retained[1] (43)	43		43		35		32		15	
Items excluded[2] (37)	37		35		29		27		13	
	%	[95% CI]	%	[95% CI]	%	[95% CI]	%	[95% CI]	%	[95% CI]
All dimensions										
(*P* +)	100	[93–100]	100	[93–100]	81	[67–92]	74	[59–86]	35	[21–51]
(*P* −)	0	[00–80]	5	[01–18]	22	[10–38]	27	[14–44]	65	[47–80]
Physical activities										
(*P* +)	100	[79–100]	100	[79–100]	81	[54–96]	81	[54–96]	37	[15–66]
(*P* −)	0	[0–6]	3	[0–11]	20	[11–32]	28	[18–41]	66	[53–77]
Mental health										
(*P* +)	100	[75–100]	100	[75–100]	69	[39–91]	69	[39–91]	46	[19–75]
(*P* −)	0	[0–5]	3	[0–10]	18	[10–29]	25	[16–37]	67	[55–78]
Pain	100	[40–100]	100	[40–100]	75	[19–99]	100	[40–100]	50	[7–93]
(*P* +)										
(*P* −)	0	[0–5]	3	[0–9]	20	[11–30]	28	[18–39]	66	[54–76]
Social support										
(*P* +)	100	[40–100]	100	[40–100]	100	[40–100]	50	[7–93]	0	[0–60]
(*P* −)	0	[0–5]	3	[0–9]	21	[13–32]	25	[16–36]	63	[51–74]
Social activities										
(*P* +)	100	[30–100]	100	[30–100]	100	[30–100]	67	[9–99]	0	[0–71]
(*P* −)	0	[0–5]	3	[0–9]	21	[12–32]	26	[17–37]	64	[52–74]

[1]Items retained in the OAKHQOL, [2]Items excluded from the 80 verbatims produced; OAKHQOL (OsteoArthritis Knee and Hip Quality-Of-Life).
P+ was defined as the number of items generated by the method that were also included in the final questionnaire over the total number of items of the final questionnaire. *P*− was defined as the total number of items not generated by the method that were also not included in the final questionnaire over the number of items not included in the final questionnaire.

Results

The initial analysis of interview transcripts identified 80 verbatim categories or potential items.

Semi-structured and cognitive interviews of patients accounted for 80 and 78 items, respectively, of the initial list, whereas focus groups of both patients and health professionals and interviews with health professionals yielded 59, 28 and 64 items, respectively.

P+, *P*− and the confidence intervals of these proportions for the various techniques are shown in Table 1. Compared with other methods, individual patient interviews produced greater *P*+ ($p < 0.01$) (with the exception of individual interviews with health professionals) and lower *P*− ($p < 0.01$), and the health professional focus group gave lower *P*+ ($p < 0.01$) and higher *P*− ($p < 0.01$) than that of all the other methods. The *P*+ and *P*− for semi-structured and cognitive patient interviews were not significantly different. The same was true for individual interviews with health professionals and patient focus groups. With regard to the physical activities domain, the health professional focus-group method produced significantly lower *P*+ than the other methods ($p < 0.01$). For the mental health domain, interviews with individual patients yielded a significantly higher *P*+ than did other methods ($p < 0.01$). For social support items, data from heath professional focus groups gave *P*+ lower than that from patients or health professional individual interviews. The other comparisons failed to show significant differences.

The proportion of patients or health professionals who provided the items eventually included in the OAKHQOL are displayed in Table 2. The two interview techniques for patients differed significantly in eliciting items concerning social support: three items of the social support dimension were more often highlighted in cognitive interviews. Significantly more patients in semi-structured interviews than

Table 2 Proportion of patient or health professional methods that provided the items that were eventually included in the OAKHQOL

Abbreviated item content of OAKHQOL (43 items)	Patient semi-structured interviews ($n = 32$)	Patient cognitive interviews ($n = 32$)	HP individual interviews ($n = 16$)	Patient focus groups ($n = 9$ and 6)[1]	HP focus group ($n = 12$)[1]
Physical activities domain					
Walking	97	84	56	100	100
Bending or straightening	66	53	31	100	100
Carrying heavy things	16	25	0	0	0
Climbing stairs	94	28	50	50	100
Going down stairs	94	44	56	100	0
Taking a bath	25	16	19	50	0
Dressing	41	16	19	0	0
Cutting toe-nails	16	3	0	0	0
Staying in the same position for a long time	69	38	6	100	100
Getting moving after staying in the same position	47	31	44	100	100
Needing a stick to walk	53	28	31	100	0
Needing help	69	31	19	100	100
Getting in and out of a car	34	16	13	100	0
Using public transport	22	16	13	100	0
Needing to spare oneself	53	31	13	100	0
Taking longer to do things	34	34	0	100	0
Pain domain					
Intensity of pain	66	47	31	100	100
Frequency of pain	41	38	0	50	0
Having difficulties getting to sleep because of pain	6	31	19	100	100
Waking up at night because of pain	44	47	6	100	0
Mental health domain					
Feeling depressed because of pain	78	75	94	100	100
Feeling older than your years	3	25	0	0	0
Being afraid of being dependent on others	13	38	0	0	100
Being afraid of becoming an invalid	78	63	69	100	100
Being embarrassed to be seen by other people	53	22	6	50	100
Worry	28	25	13	0	100
Feeling depressed	22	66	0	0	0
Wondering what is going to happen	22	41	0	50	0
Worry about the side-effects of treatment	44	50	38	100	100
Impaired family life	34	9	13	100	0

(*continued*)

Table 2 *Continued*

Abbreviated item content of OAKHQOL (43 items)	Patient semi-structured interviews ($n = 32$)	Patient cognitive interviews ($n = 32$)	HP individual interviews ($n = 16$)	Patient focus groups ($n = 9$ and 6)[1]	HP focus group ($n = 12$)[1]
Feeling aggressive and irritable	34	13	19	50	0
Feelings of being a burden to close relatives	34	19	13	50	0
Feeling embarrassed to ask for help	47	6	19	100	0
Social support domain					
Talking about arthritis-related problems	28	78	25	100	0
Feeling that others understand arthritis-related problems	28	47	31	100	0
Feeling supported by people close to me	22	63	6	0	0
Feeling supported by people around	9	38	6	0	0
Social functioning domain					
Ability to plan for the future	28	59	19	0	0
Going out whenever one likes	50	50	44	100	0
Having friends in whenever one likes	19	31	44	100	0
Independent items					
Impaired professional activity	22	34	44	50	0
Impaired life with partner	13	31	13	100	0
Restricted sexual life	9	9	6	0	100

HP, health professionals; Values are percentages.
[1]Percentages are 0, 50 or 100% because the unit is the group and not the patient. Abbreviated item content of the OAKHQOL (OsteoArthritis Knee and Hip Quality-of-Life) are not the whole items; the translation adaptation in English is in progress.

in cognitive interviews talked about physical activities domain items: difficulties with stairs, getting dressed, remaining in the same position and needing help or a stick to walk. With regard to the other domains of the OAKHQOL, the two methods varied in the frequency with which they elicited various items. The items 'being embarrassed to be seen by other people', 'feeling embarrassed to ask for help' and 'feelings of being a burden to close relatives' were more commonly mentioned in semi-structured interviews. Cognitive interviews were more likely to include items concerning fear of future dependency, feeling older than one's age and perspective on life.

Individual health professional interviews frequently elicited typical functional disability items: difficulties with walking, getting up or down stairs, or moving after staying in the same position. Apart from reporting these three items, health professionals were less likely (<20%) to report mental health items.

Discussion

The use of several methods to generate items during the development of the OAKHQOL allowed us to study the relative contributions of the methods in obtaining oral material. The different methods were not equivalent in generating items and clearly did not produce the same results. Individual interviews produced a greater proportion of items and more verbatim remarks compared to focus groups, and patients contributed more items than health professionals.

One recognized advantage of focus groups is that interactions between participants may reveal specific themes or

new ideas. However, here, no resulting item was identified only by focus groups.

Health professionals provided no ideas or remarks not suggested by patients. Individual interviews with health professionals provided the principal functional disability items, but few health professionals emphasized everyday physical activities such as dressing, bathing and getting out of a car. With the exception of three frequently specified items, they rarely reported mental health items. They also rarely described support from others as a critical concern for patients. In accordance with other studies, the differences between patients and health professionals appear more pronounced in the physical, psychological and activity domains than in the social domain [15, 17]. Moreover, professionals often provided items similar to those included in other scales used in osteoarthritis [30–32].

Cognitive interviews to generate quality-of-life items have rarely been used. A comparison of patient interviews with cognitive or semi-structured techniques revealed the former technique providing more detailed items for exchanges with other people, social support, perspective on life and fear of dependency. Semi-structured interviews produced more physical activity items and items concerning embarrassment at being seen and concerns about asking for help and being a burden to close relatives.

Accurate and relevant outcome measures to adequately capture aspects of life most important to people living with a chronic disease are needed to improve patient-centered care. To better understand what matters to patients, their concerns and priorities for care, outcome measures should be carefully developed. In this regard, the item-generation step is critical to ensure good content of the instrument. One strength of this study is that a large and relevant group of patients and experts helped generate items. The range of interview and focus-group methods was also extensive, and several analysts were consulted. The resulting questionnaire allowed for a broad and deep exploration of the concerns, needs and values of osteoarthritis patients.

However, some aspects of the study design might have affected the results. First, in the absence of guidelines or a 'standard' for 'good practice' in composite scale construction, the final OAKHQOL was used as the reference to compare item-generation methods. Obviously, it was created with the methods themselves, and the large number of items provided by individual patient interviews increased the likelihood of their presence in the final version of the questionnaire.

Also, groups of interviewees differed in size, and patients had more opportunity to express themselves than did health professionals. The number of individual patient interviews was twice that of health professional interviews, and the same imbalance was true for the focus groups. However, the proportion of health professionals interviewed individually who evoked each of the 43 items was much lower than that of patients interviewed (also, we probably could have reached 'saturation' of information with fewer than 32 individual patient interviews). The focus groups were approximately one-quarter the size of the number of subjects individually interviewed, which decreased the opportunity for focus groups to reach saturation of information.

In addition, the development of a guide for use in individual patient interviews but not for the other methods could have allowed for generation of more items with the former method, with their better chance of being included in the final questionnaire.

As well, the efficiency of focus groups and individual interviews might depend on the length of time participants are given to speak. Participants in focus groups clearly have less opportunity to speak, which is inherent to the method; indeed, focus groups aim to allow those taking part to explore what they have in common, rather than put forward their individual points of view.

Finally, despite constant efforts to maintain objectivity, the panel of experts might have impressed their view of the quality-of-life concept on the constructed instrument.

The final OAKHQOL was used as the reference to compare item-generation methods, but other references could have been chosen, such as a selection of items based on item-response theory. However, the methodological approach to shortening our composite measurement scale followed published recommendations and can be considered as a reference [27].

In conclusion, this work highlights the need to structure and report how items are generated during the construction phase of an instrument. The different methods used to elicit oral material are not equivalent. Themes are not necessarily revealed by all methods, and some themes are more likely to emerge when one method is used rather than another. Individual interviews with patients identified all 80 items produced during development of the OAKHQOL and all the 43 items included in the final version. The other methods contributed no additional information.

We recommend that (i) when a new scale is published, the generation of items should be formalized and reported; (ii) details should include the methods used and the number of patients or health professionals involved; (iii) individual interviewing with patients is the best method for generating quality-of-life instruments; and (iv) with regard to individual patients, cognitive and semi-structured interviews may be combined because they are complementary. These recommendations are provisional, and other studies are necessary to confirm the present results.

Acknowledgments

M. Alfonsi, F. Arnould, X. Ayral, P. Bouillot, I. Cahitte, I. Chary-Valckenaere, P. Chaspoux, C. Cuny, J.P. Delagoutte, D. Dumont-Fisher, P. Fener, H. Gaudin, M.F. Gérard, F. Goupy Le Maitre, J.S. Lequintrec, J.P. Gros, D. Hannouche, E. Hiraux, N. Jeanson, J.M. Lardry, D. Loeuille, D. Mainard, D. Nebout, G. Osnowycz, H. Ouakil, S. Perrot, S. Poiraudeau, J. Pourel, P. Prost, P. Rabany, M. Rousseau, P. Thomas, F. Touzard, G. Vançon, L. Vastel, J.P. Voilquin.

This study was supported by the Clinical Epidemiology Center, INSERM (Institut National de la Santé et de

la Recherche Médicale) CIE 6-Health Ministry, Nancy University Hospital.

Funding

The French Quality of Life in Rheumatology Group.

References

1. Steiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to their Development and Use*. Oxford: Oxford University Press, 1995, p. 231.
2. Coste J, Fermanian J, Venot A. Methodological and statistical problems in the construction of composite measurement scales: a survey of six medical and epidemiological journals. *Stat Med* 1995;**14**:331–45.
3. Shaw MJ, Talley NJ, Beebe TJ *et al*. Initial validation of a diagnostic questionnaire for gastroesophageal reflux disease. *Am J Gastroenterol* 2001;**96**:52–7.
4. Grant EN, Turner-Roan K, Daugherty SR *et al*. Development of a survey of asthma knowledge, attitudes, and perceptions: the Chicago Community Asthma Survey. Chicago Asthma Surveillance Initiative Project Team. *Chest* 1999;**116**:178S–83S.
5. Niero M, Martin M, Finger T *et al*. A new approach to multicultural item generation in the development of two obesity-specific measures: the Obesity and Weight Loss Quality of Life (OWLQOL) questionnaire and the Weight-Related Symptom Measure (WRSM). *Clin Ther* 2002;**24**:690–700.
6. Harris-Kojetin LD, Fowler FJ, Jr, Brown JA *et al*. The use of cognitive testing to develop and evaluate CAHPS 1.0 core survey items. Consumer Assessment of Health Plans Study. *Med Care* 1999;**37**:MS10–21.
7. Paterson C, Britten N. Acupuncture for people with chronic illness: combining qualitative and quantitative outcome assessment. *J Altern Complement Med* 2003;**9**:671–81.
8. Drennan J. Cognitive interviewing: verbal data in the design and pretesting of questionnaires. *J Adv Nurs* 2003;**42**:57–63.
9. Collins D. Pretesting survey instruments: an overview of cognitive methods. *Qual Life Res* 2003;**12**:229–38.
10. Fisher RP, Falkner KL, Trevisan M *et al*. Adapting the cognitive interview to enhance long-term (35 years) recall of physical activities. *J Appl Psychol* 2000;**85**:180–9.
11. Jobe JB. Cognitive psychology and self-reports: models and methods. *Qual Life Res* 2003;**12**:219–27.
12. Geiselman RE, Fisher RP, MacKinnon DP *et al*. Eyewitness memory enhancement in the police interview: cognitive retrieval mnemonics versus hypnosis. *J Appl Psychol* 1985;**70**:401–12.
13. Kitzinger J. Qualitative research. Introducing focus groups. *BMJ* 1995;**311**:299–302.
14. Dilorio C, Hockenberry-Eaton M, Maibach E *et al*. Focus groups: an interview method for nursing research. *J Neurosci Nurs* 1994;**26**:175–80.
15. Woodend AK, Nair RC, Tang AS. Definition of life quality from a patient versus health care professional perspective. *Int J Rehabil Res* 1997;**20**:71–80.
16. Rothwell PM, McDowell Z, Wong CK *et al*. Doctors and patients don't agree: cross sectional study of patients' and doctors' perceptions and assessments of disability in multiple sclerosis. *BMJ* 1997;**314**:1580–3.
17. Lomas J, Pickard L, Mohide A. Patient versus clinician item generation for quality-of-life measures. The case of language-disabled adults. *Med Care* 1987;**25**:764–9.
18. McHorney CA, Bricker DE, Kramer AE *et al*. The SWAL-QOL outcomes tool for oropharyngeal dysphagia in adults: I. Conceptual foundation and item development. *Dysphagia* 2000;**15**:115–21.
19. Lo IK, Griffin S, Kirkley A. The development of a disease-specific quality of life measurement tool for osteoarthritis of the shoulder: The Western Ontario Osteoarthritis of the Shoulder (WOOS) index. *Osteoarthritis Cartilage* 2001;**9**:771–8.
20. Rouland JF, Denis P, Bechetoille A *et al*. [Creating a specific quality-of-life questionnaire in patients with glaucoma: item generation]. *J Fr Ophtalmol* 2002;**25**:785–94.
21. Gee L, Abbott J, Conway SP *et al*. Development of a disease specific health related quality of life measure for adults and adolescents with cystic fibrosis. *Thorax* 2000;**55**:946–54.
22. Rat AC, Coste J, Pouchot J *et al*. OAKHQOL: a new instrument to measure quality of life in knee and hip osteoarthritis. *J Clin Epidemiol* 2005;**58**:47–55.
23. Guillemin F, Coste J, Retel-Rude N *et al*. Strategy and methodology for choice of items in psychometric measurement: designing a quality of life instrument for of hip and knee osteoarthritis. In: Mesbah M, Cole BF, Ting Lee ML (eds). *Statistical Methods for Quality of Life Studies: Design, Measurements and Analysis*. Dordrecht: Kluwer Academic Publishers, 2002, 51–61.
24. Morgan DL. (ed) *Focus Groups as Qualitative Research*. vol. 16, 2nd edn. Thousand Oaks, CA: Sage, 1997.
25. Altman R, Alarcon G, Appelrouth D *et al*. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheum* 1991;**34**:505–14.
26. Altman R, Asch E, Bloch D *et al*. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. *Arthritis Rheum* 1986;**29**:1039–49.
27. Coste J, Guillemin F, Pouchot J *et al*. Methodological approaches to shortening composite measurement scales. *J Clin Epidemiol* 1997;**50**:247–52.
28. Guillemin F, Coste J, Pouchot J *et al*. The AIMS2-SF: a short form of the Arthritis Impact Measurement Scales 2. French Quality of Life in Rheumatology Group. *Arthritis Rheum*. 1997;**40**:1267–74.
29. World Health Organization. The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. *Soc Sci Med* 1995;**41**:1403–9.

30. Bellamy N, Buchanan WW, Goldsmith CH *et al*. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988;**15**:1833–40.

31. Lequesne MG, Mery C, Samson M *et al*. Indexes of severity for osteoarthritis of the hip and knee. Validation–value in comparison with other assessment tests. *Scand J Rheumatol Suppl* 1987;**65**:85–9.

32. Ware JE, Jr, Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. *J Clin Epidemiol* 1998;**51**:903–12.

Accepted for publication 8 August 2007